IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application : **10/596135**

Applicant(s) : **PAUWS, Steffen Clarence**

Filed : **6/1/2006**

Confirmation : **2675**

T.C./Art Unit : **2169**

Examiner : **VO, Cecile H.**

Atty. Docket : **NL031435US**

Title: **SEARCHING IN A MELODY DATABASE**

Mail Stop: **APPEAL BRIEF - PATENTS**
Commissioner for Patents
Alexandria, VA 22313-1450

APPEAL UNDER 37 CFR 41.37

Sir:

This is an appeal from the decision of the Examiner dated 18 July 2008, finally rejecting claims 1-20 of the subject application.

This paper includes (each beginning on a separate sheet):

1. **Appeal Brief;**
2. **Claims Appendix;**
3. **Evidence Appendix; and**
4. **Related Proceedings Appendix.**

APPEAL BRIEF

I. REAL PARTY IN INTEREST

The above-identified application is assigned, in its entirety, to **Koninklijke Philips Electronics N. V.**

II. RELATED APPEALS AND INTERFERENCES

Appellant is not aware of any co-pending appeal or interference that will directly affect, or be directly affected by, or have any bearing on, the Board's decision in the pending appeal.

III. STATUS OF CLAIMS

Claims 1-20 are pending in the application.

Claims 1-20 stand rejected by the Examiner under 35 U.S.C. 102(e).

These rejected claims are the subject of this appeal.

IV. STATUS OF AMENDMENTS

No amendments were filed subsequent to the final rejection in the Office Action dated 18 July 2008. A reply to the final rejection was filed on 16 September 2008.

V. SUMMARY OF CLAIMED SUBJECT MATTER

This invention addresses a method and system for searching a melody database for a melody that best matches a sample audio fragment, such as a sample of a user's singing, humming, whistling, tapping, and so on (Applicant's specification, page 1, lines 2-5, 19-20). The inventor has recognized that a user's recollection of a song is often spotty, or patchy, and that a disjoint input will generally be counter-productive for accurate searching (page 2, lines 5-14). Although conventional melody search engines may allow for relatively minor gaps or timing-errors in the audio fragment, the algorithms used for processing melodies naturally presume that the elements/notes in the audio fragment are in proper time-sequence order (page 2, lines 11-12). In embodiments of the applicant's invention, the audio fragment is partitioned into a plurality of query sub-strings, each sub-string is independently used to search the melody database for melodies that best match the sub-string, and the best match to the audio fragment is based on these sub-string searches (page 2, lines 14-17). For example, a song that closely matches all of the sub-strings is scored higher than one that only matches a few. In this manner, each sub-string can be processed as a time-ordered sequences of notes, but the set of sub-strings are not presumed to be ordered in time (page 2, lines 16-20). Preferably, the sub-strings correspond to phrases of the melody, and techniques are provided for identifying phrase boundaries; for example, if the input audio segment is multi-modal, each modality change is assumed to signal a start of a new phrase (page 2, line 30 – page 3, line 2; page 3, lines 12-13).

As claimed in independent claim 1, the invention comprises a method (FIGs. 1 and 2) comprising:

decomposing (117) a query string that corresponds to an encoding of an audio fragment into a sequence of a plurality of query sub-strings (page 8, line 25; FIG. 3);

independently searching (118) a melody database for at least a respective closest match for each sub-string of the plurality of query sub-strings (page 8, lines 25-26); and

in dependence on search results for the respective sub-strings, determining (119) at least a closest match for the query string (page 8, lines 26-27).

As claimed in dependent claim 2, the invention comprises the method of claim 1, wherein decomposing the query string includes decomposing the query string into sub-strings that each substantially corresponds to a phrase of a melody (page 2, lines 30-31).

As claimed in dependent claim 5, the invention comprises the method of claim 3, wherein the query string includes a plurality of query input modalities and a change in query input modality substantially coincides with a sub-string boundary (page 3, lines 12-13).

As claimed in dependent claim 6, the invention comprises the method of claim 1, wherein decomposing the query string includes (FIG. 3):

estimating (310) how many (Ns) sub-strings are present in the query string (page 9, lines 8-10);

dividing (320) the query string in Ns sequential sub-strings, each sub-string being associated with a respective centroid that represents the sub-string (page 9, lines 14-15; 19-20);

iteratively (330-350):

for each centroid, determining a respective centroid value in dependence on the sub-string associated with the respective centroid (page 9, line 33 – page 10, line 2); and

determining, for each of the sub-strings, corresponding sub-string boundaries by minimizing (330) a total distance measure between each of the centroids and the sub-string associated with the respective centroid (page 10, lines 2-3);

until (350) a predetermined convergence criterion is met (page 10, lines 13-14).

As claimed in dependent claim 7, the invention comprises the method of claim 6, wherein estimating how many (Ns) sub-strings are present in the query string includes dividing a duration of the audio fragment by an average duration of a phrase (page 9, lines 15-18).

As claimed in dependent claim 8, the invention comprises the method of claim 5, wherein decomposing the query string includes retrieving for each of the input modalities a respective classification criterion and detecting the change in query input modality based on the classification criteria (page 11, lines 14-18).

As claimed in dependent claim 10, the invention comprises the method of claim 1, wherein searching for each sub-string in the database includes generating for the sub-string an N-best list (N >=2) of the N closest corresponding parts in the database with a corresponding measure of resemblance (page 8, lines 6-8); and performing the determining of the at least closest match for the query string based on the measures of resemblance of the N-best lists of the sub-strings (page 8, lines 13-15).

As claimed in independent claim 11, the invention comprises a computer media that includes a computer program product (page 13, lines 29-33) operative to cause a processor to (FIGs. 1 and 2):

decompose (117) a query string that corresponds to an encoding of an audio fragment into a sequence of a plurality of query sub-strings (page 8, line 25);

independently search (118) a melody database for at least a respective closest match for each sub-string of the plurality of query sub-strings (page 8, lines 25-26); and

in dependence on the search results for the respective sub-strings, determine (119) at least a closest match for the query string (page 8, lines 26-27).

As claimed in independent claim 12, the invention comprises a system (FIGs. 1 and 2) comprising:

an input (122) for receiving a query string that corresponds to an encoding of an audio fragment from a user (page 6, lines 14-15);

a melody database (114) for storing respective representations of plurality of audio fragments (page 7, lines 2-4);

at least one processor (116) that is configured to:

decompose (117) the query string into a sequence of a plurality of query sub-strings (page 8, line 25);

search (118) the database for at least a respective closest match for each sub-string of the plurality of query sub-strings (page 8, lines 25-26); and

determine (119) at least a closest match for the query string based on the closest matches for the plurality of query sub-strings (page 8, lines 26-27).

As claimed in dependent claim 13, the invention comprises the system of claim 12, wherein each sub-string substantially corresponds to a phrase of a melody (page 2, lines 30-31).

As claimed in dependent claim 16, the invention comprises the system of claim 14, wherein the query string includes a plurality of query input modalities, and a change in query input modality substantially coincides with a sub-string boundary (page 3, lines 12-13).

As claimed in dependent claim 17, the invention comprises the system of claim 16, wherein the processor is configured to decompose the query string by:

retrieving for each of the input modalities a respective classification criterion (page 11, lines 14-17) and

detecting the change in query input modality based on the classification criteria (page 11, lines 17-18).

As claimed in dependent claim 18, the invention comprises the system of claim 12, wherein the processor is configured to decompose the query string by (FIG. 3):

estimating (310) how many (Ns) sub-strings are present in the query string (page 9, lines 8-10);

dividing (320) the query string in Ns sequential sub-strings (page 9, lines 14-15); each sub-string being associated with a respective centroid that represents the sub-string (page 9, lines 19-20);

iteratively (330-350):

for each centroid, determining a respective centroid value in dependence on the sub-string associated with the respective centroid (page 9, line 33 – page 10, line 2); and

determining, for each of the sub-strings, corresponding sub-string boundaries by minimizing (330) a total distance measure between each of the centroids and the sub-string associated with the respective centroid (page 10, lines 2-3);

until (350) a predetermined convergence criterion is met (page 10, lines 13-14).

As claimed in dependent claim 19, the invention comprises the system of claim 18, wherein estimating how many (Ns) sub-strings are present in the query string includes dividing a duration of the audio fragment by an average duration of a phrase (page 9, lines 15-18).

As claimed in dependent claim 20, the invention comprises the system of claim 12 wherein the at least one processor is configured to generate for each sub-string an N-best list (N >=2) of the N closest corresponding parts in the database with a corresponding measure of resemblance (page 8, lines 6-8), and determine the at least closest match for the query string based on the measures of resemblance of the N-best lists of the sub-strings (page 8, lines 13-15).

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

Claims 1-20 stand rejected under 35 U.S.C. 102(e) over Tsui et al. (USPA 2007/0163425, hereinafter Tsui).

VII. ARGUMENT

Claims 1-20 stand rejected under 35 U.S.C. 102(e) over Tsui

Claims 1-20

Tsui fails to teach independently searching a melody database for at least a respective closest match for each sub-string of a plurality of query sub-strings, and fails to teach determining at least a closest match for the query string based on the search results for such sub-strings, as specifically claimed in each of the applicant's independent claims 1, 11, and 12.

The Office action asserts that Tsui teaches independently searching a melody database for at least a respective closest match for each sub-string of a plurality of query sub-strings at paragraph 0044, lines 1-4. The applicant respectfully disagrees with this assertion.

Tsui teaches that each song or piece of music is converted into note and timing data that is stored in a music database 14. To find a song matching an input query, the query is also converted into note and timing data 150, and a note matching engine 16 searches a music database 14 for a match to this note and timing data 150 representing the query.

The Office action maintains that each note of the input query corresponds to a sub-string. However, the applicant notes that Tsui does not teach independently searching the database 14 for each note of the query, as the Office action's proposed interpretation of Tsui requires in order to read upon the applicant's claims. At the cited text, Tsui teaches:

> "The note matching engine 16 compares the differential note and timing file 150 from the melody-to-note conversion subsystem 12 with songs or pieces of music in the music reference database 14, which are stored in a similar file format." (Tsui, 0044, lines 1-4.)

As is clearly evident, the cited text does not disclose searching a database for each of a plurality of sub-strings of a query string, and specifically does not teach independently searching the database for each note of a query string, as asserted in the Office action.

Because Tsui does not teach independently searching a database for a closest match to each sub-string of a query, Tsui cannot be said to teach determining a closest match for the query string based on the search results for such sub-strings. The Office action asserts that Tsui provides this teaching at paragraph 0044, lines 20-21, paragraph 0045. At the cited text, Tsui teaches:

> "The engine 16 calculates a matching score for each song in the database 14.
> The output subsystem 18 sorts the songs or music in the database 16 based on the matching scores. The highest ranked song(s) or piece(es) of music is selected for presentation to the user." (Tsui, 0044, lines 20-21; 0045.)

As is clearly evident, at the cited text, Tsui teaches determining a single matching score for the query string's matching to each song; Tsui does not teach determining a matching score for each sub-string of the query string. Tsui's highest-scoring song is based directly upon the single score associated with each song, and is not based upon the plurality of sub-string matching scores determined for each song.

The applicant further note that the Office action's proposed use of Tsui's teaching would render Tsui's device unsuitable for its intended function of finding a match to a query string. The Office action asserts that Tsui's device could be used to independently find a matching score for each note in a query string, and then determine the closest matching song based on these independent note-matching scores. The applicant respectfully notes that songs are characterized by a sequence of notes; the mere presence of the same individual notes in two pieces of music does not indicate, per se, a similarity between the two pieces. If one song includes a sequence of notes A-B-C-D-E, and another includes a sequence D-B-A-E-C, these two song sequences would not be considered similar; to the contrary, they would be considered significantly different, even though they contain the same individual notes.

In the Office action's proposed interpretation of Tsui, if the query string includes the note "C", the database would be searched to find the song that most closely matches this query note/sub-string. The applicant respectfully maintains that the statement of the task itself, "find a song that most closely matches a query-note", is virtually meaningless; does a song most closely match the query-note "C" because it contains more "C" notes than other songs? If another query-string included two "C" notes, and the database is independently searched for each note, would the matching score for each song be doubled?

Because Tsui fails to teach independently searching a melody database for at least a respective closest match for each sub-string of a plurality of query sub-strings, and fails to teach determining at least a closest match for the query string based on the search results for such sub-strings, and because the Office action's proposed interpretation of Tsui would not provide a viable melody search technique, the applicant respectfully maintains that the rejection of claims 1-20 under 35 U.S.C. 102(e) over Tsui is unfounded, and should be reversed by the Board.

Claims 2 and 13

Tsui fails to teach decomposing a query string into sub-strings that each substantially corresponds to a phrase of a melody, as specifically claimed in each of claims 2 and 13. As noted above, in the rejection of claims 1 and 12, upon which claims 2 and 13 depend, the Office action maintains that each note in the query string corresponds to the claimed query sub-string. The applicant respectfully maintains that this asserted correspondence is contrary to an assertion that the query sub-strings correspond to a phrase of a melody.

The Office action asserts that Tsui teaches decomposing a query string into sub-strings that each substantially correspond to a phrase of a melody at paragraph 0042, lines 1-4. The applicant respectfully disagrees with this assertion. At the cited text, Tsui teaches:

> "The melody-to-note conversion subsystem 12 converts the digitized input melody 20 into a sequence of musical notes characterized by pitch, beat duration and confidence levels." (Tsui, 0042, lines 1-4.)

As is clearly evident, the above cited text does not disclose decomposing the query string into sub-strings corresponding to phrases of a melody, as asserted in the Office action.

Because Tsui fails to teach each of the elements of claims 2 and 13, the applicant respectfully maintains that the rejection of claims 2 and 13 under 35 U.S.C. 102(e) over Tsui is unfounded, and should be reversed by the Board.

Claims 5, 8, and 16-17

Tsui fails to teach that a change in query input modality substantially coincides with a sub-string boundary, as specifically claimed in each of claims 5 and 16, upon which claims 6 and 17 depend.

The Office action asserts that Tsui provides this teaching at paragraph 0048. The applicant respectfully disagrees with this assertion. At the cited text, Tsui teaches:

> "a list of breakpoints, which indicate the boundaries between distinct notes in the input melody" (Tsui, 0048.)

As is clearly evident, the above cited text does not address changes of input modality, and specifically does not disclose that a change in query input modality substantially coincides with a sub-string boundary, as asserted in the Office action.

Because Tsui fails to teach each of the elements of claims 5 and 16, the applicant respectfully maintains that the rejection of claims 5, 8, 16, and 17 under 35 U.S.C. 102(e) over Tsui is unfounded, and should be reversed by the Board.

Claims 8 and 17

Tsui fails to teach detecting the change in query input modality based on a classification criteria of each input modality, as specifically claimed in claims 8 and 17.

The Office action asserts that Tsui provides this teaching at paragraph 104. The applicant respectfully disagrees with this assertion. At the cited text, Tsui teaches an alternative technique for determining spectral energy distribution (SED), and does not address classifying criteria for each input modality, as asserted in the Office action.

Because Tsui fails to teach each of the elements of claims 8 and 17, the applicant respectfully maintains that the rejection of claims 8 and 17 under 35 U.S.C. 102(e) over Tsui is unfounded, and should be reversed by the Board.

Claims 6-7 and 18-19

Tsui fails to teach estimating how many (Ns) sub-strings are present in the query string; fails to teach dividing the query string in Ns sequential sub-strings, each sub-string being associated with a respective centroid that represents the sub-string; and fails to teach iteratively determining a respective centroid value in dependence on the sub-string, and determining the sub-string boundaries by minimizing a total distance measure between each of the centroids and the corresponding sub-string, until a predetermined convergence criterion is met, as specifically claimed in claims 6 and 18, upon which claims 7 and 19 depend.

The Office action asserts that Tsui teaches determining the sub-string boundaries by minimizing a total distance measure between each of the centroids and corresponding sub-string at paragraph 0011. The applicant respectfully disagrees with this assertion. At the cited text, Tsui teaches:

> "One aspect of the invention provides a method and related system for converting a digitized melody into a sequence of notes. Generally speaking, the method involves estimating breakpoints in the input melody based on changes in the distribution of energy across the frequency spectrum over time. In the preferred embodiment, the melody is segmented into a series of frames. A spectral energy distribution (SED) indicator is computed for each frame and at least initial breakpoints estimates are derived based on the SED indicator. Notes are defined between adjacent breakpoints." (Tsui, 0011.)

As is clearly evident, the above cited text does not address determining distance measures, and specifically does not disclose determining the sub-string boundaries by minimizing a total distance measure between each centroid and corresponding sub-string, as asserted in the Office action.

Further, the Office action asserts that Tsui teaches performing this minimizing iteratively until a predetermined convergence criterion is met at paragraph 0008. The applicant respectfully disagrees with this assertion. At the cited text, Tsui teaches:

> "One aspect of the invention provides a method and system for converting a digitized melody into a series of notes. The method and system receive a digitized representation of an input melody, identify breakpoints in the melody in order to define notes therein, determine a pitch and beat duration for each note of the melody, and associate a confidence level with each breakpoint, or each note, or both." (Tsui, 0008.)

As is clearly evident, the above cited text does not address performing an iterative process, and specifically does not disclose iteratively minimizing a total distance measure between each of the centroids and the corresponding sub-string until a predetermined convergence criterion is met, as asserted in the Office action.

Because Tsui fails to teach each of the elements of claims 6 and 18, the applicant respectfully maintains that the rejection of claims 6-7 and 18-19 under 35 U.S.C. 102(e) over Tsui is unfounded, and should be reversed by the Board.

Claims 7 and 19

Tsui fails to teach dividing a duration of the audio fragment by an average duration of a phrase to estimate how many sub-strings are present, as specifically claimed in claims 7 and 19.

The Office action asserts that Tsui provides this teaching at paragraph 0010. The applicant respectfully disagrees with this assertion. At the cited text, Tsui teaches:

> "In the preferred embodiment, segmentation of the input melody into distinct notes divided by breakpoints is based on changes or differences in the distribution of energy across the frequency spectrum over time. The confidence levels associated with each breakpoint and/or note may be based on changes in pitch, as well as absolute and relative values of a spectral energy distribution indicator." (Tsui, 0010.)

As is clearly evident, the above cited text discloses that the note boundaries are determined based on changes in energy levels, and does not address estimating how many sub-strings are present based on an average duration of a phrase, as asserted in the Office action.

Because Tsui fails to teach each of the elements of claims 7 and 19, the applicant respectfully maintains that the rejection of claims 7 and 19 under 35 U.S.C. 102(e) over Tsui is unfounded, and should be reversed by the Board.

Claim 10 and 20

Tsui fails to teach generating a list of the N closest corresponding parts in the database for each sub-string with a corresponding measure of resemblance; and fails to teach determining the at least closest match for the query string based on the measures of resemblance of the N-best lists of the sub-strings, as specifically claimed in each of claims 10 and 20.

The Office action asserts that Tsui teaches generating a list of the N closest corresponding parts in the database for each sub-string with a corresponding measure of resemblance at paragraph 0048. The applicant respectfully disagrees with this assertion. At the cited text, Tsui teaches:

> "a list of breakpoints, which indicate the boundaries between distinct notes in the input melody" (Tsui, 0048.)

As is clearly evident, the above cited text fails to address a list of closest corresponding parts for each substring, and specifically fails to disclose generating a list of the N closest corresponding parts in the database for each sub-string with a corresponding measure of resemblance, as asserted in the Office action.

The Office action further asserts that Tsui teaches determining the closest match for the query string based on the measures of resemblance of the N-best lists of the sub-strings at paragraphs 0050-0051. The applicant respectfully disagrees with this assertion. At the cited text, Tsui teaches techniques for determining breakpoints between notes, and associating a confidence level to the breakpoints and/or the corresponding notes. The cited text details a process that is performed during the partitioning of a query into individual notes, prior to the process of comparing the query to pieces of music in the database, and thus cannot be said to teach determining a closest match for the query based on matches for the sub-strings, as asserted in the Office action.

Because Tsui fails to teach each of the elements of claims 10 and 20, the applicant respectfully maintains that the rejection of claims 10 and 20 under 35 U.S.C. 102(e) over Tsui is unfounded, and should be reversed by the Board.

CONCLUSIONS

Because Tsui fails to teach independently searching a melody database for at least a respective closest match for each sub-string of a plurality of query sub-strings, and fails to teach determining at least a closest match for the query string based on the search results for such sub-strings, and because the Office action's proposed interpretation of Tsui would not provide a viable melody search technique, the Applicant respectfully requests that the Examiner's rejection of claims 1-20 under 35 U.S.C. 102(e) be reversed by the Board, and the claims be allowed to pass to issue.

Because Tsui fails to teach the elements of the dependent claims discussed above, the Applicant respectfully requests that the Examiner's rejection of each of claims 2, 5-8, 10, 13, and 16-20 under 35 U.S.C. 102(e) be reversed by the Board, and the claims be allowed to pass to issue.

Respectfully submitted

/Robert M. McDermott/
Robert M. McDermott, Attorney
Registration Number 41,508
patents@lawyer.com
804-493-0707

Please direct all correspondence to:
Corporate Counsel
U.S. PHILIPS CORPORATION
P.O. Box 3001
Briarcliff Manor, NY 10510-8001

CLAIMS APPENDIX

1. A method comprising:

decomposing a query string that corresponds to an encoding of an audio fragment into a sequence of a plurality of query sub-strings;

independently searching a melody database for at least a respective closest match for each sub-string of the plurality of query sub-strings; and

in dependence on search results for the respective sub-strings, determining at least a closest match for the query string.

2. The method of claim 1, wherein decomposing the query string includes decomposing the query string into sub-strings that each substantially correspond to a phrase of a melody.

3. The method of claim 1, including enabling a user to input the query string.

4. The method of claim 3, wherein the query string includes a plurality of query input modalities that includes at least one of: humming, singing, whistling, tapping, clapping, percussive vocal sounds.

5. The method of claim 3, wherein the query string includes a plurality of query input modalities and a change in query input modality substantially coincides with a sub-string boundary.

6. The method of claim 1, wherein decomposing the query string includes:

estimating how many (Ns) sub-strings are present in the query string;

dividing the query string in Ns sequential sub-strings; each sub-string being associated with a respective centroid that represents the sub-string;

iteratively:

for each centroid, determining a respective centroid value in dependence on the sub-string associated with the respective centroid; and

determining, for each of the sub-strings, corresponding sub-string boundaries by minimizing a total distance measure between each of the centroids and the sub-string associated with the respective centroid;

until a predetermined convergence criterion is met.

7. The method of claim 6, wherein estimating how many (Ns) sub-strings are present in the query string includes dividing a duration of the audio fragment by an average duration of a phrase.

8. The method of claim 5, wherein decomposing the query string includes retrieving for each of the input modalities a respective classification criterion and detecting the change in query input modality based on the classification criteria.

9. The method of claim 3, including constraining a sub-string to fall within two successive changes in query input modality.

10. The method of claim 1, wherein searching for each sub-string in the database includes generating for the sub-string an N-best list (N >=2) of the N most closest corresponding parts in the database with a corresponding measure of resemblance; and performing the determining of the at least closest match for the query string based on the measures of resemblance of the N-best lists of the sub-strings.

11. A computer media that includes a computer program product operative to cause a processor to:

decompose a query string that corresponds to an encoding of an audio fragment into a sequence of a plurality of query sub-strings;

independently search a melody database for at least a respective closest match for each sub-string of the plurality of query sub-strings; and

in dependence on the search results for the respective sub-strings, determine at least a closest match for the query string.

12. A system comprising:

an input for receiving a query string that corresponds to an encoding of an audio fragment from a user;

a melody database for storing respective representations of plurality of audio fragments;

at least one processor that is configured to:

decompose the query string into a sequence of a plurality of query sub-strings;

search the database for at least a respective closest match for each sub-string of the plurality of query sub-strings; and

determine at least a closest match for the query string based on the closest matches for the plurality of query sub-strings.

13. The system of claim 12, wherein each sub-string substantially corresponds to a phrase of a melody.

14. The system of claim 12, wherein the at least one processor is configured to enable a user to input the query string.

15. The system of claim 14, wherein the query string includes at least one of a plurality of query input modalities that includes at least one of: humming, singing, whistling, tapping, clapping, and percussive vocal sounds.

16. The system of claim 14, wherein the query string includes a plurality of query input modalities, and a change in query input modality substantially coincides with a sub-string boundary.

17. The system of claim 16, wherein the processor is configured to decompose the query string by:

retrieving for each of the input modalities a respective classification criterion and

detecting the change in query input modality based on the classification criteria.

18. The system of claim 12, wherein the processor is configured to decompose the query string by:

estimating how many (Ns) sub-strings are present in the query string;

dividing the query string in Ns sequential sub-strings; each sub-string being associated with a respective centroid that represents the sub-string;

iteratively:

for each centroid, determining a respective centroid value in dependence on the sub-string associated with the respective centroid; and

determining, for each of the sub-strings, corresponding sub-string boundaries by minimizing a total distance measure between each of the centroids and the sub-string associated with the respective centroid;

until a predetermined convergence criterion is met.

19. The system of claim 18, wherein estimating how many (Ns) sub-strings are present in the query string includes dividing a duration of the audio fragment by an average duration of a phrase.

20. The system of claim 12 wherein the at least one processor is configured to generate for each sub-string an N-best list (N >=2) of the N closest corresponding parts in the database with a corresponding measure of resemblance, and determine the at least closest match for the query string based on the measures of resemblance of the N-best lists of the sub-strings.

EVIDENCE APPENDIX

No evidence has been submitted that is relied upon by the appellant in this appeal.

RELATED PROCEEDINGS APPENDIX

Appellant is not aware of any co-pending appeal or interference which will directly affect or be directly affected by or have any bearing on the Board's decision in the pending appeal.